STANFORD RESEARCH INSTITUTE

MENLO PARK, CALIFORNIA 94025

November 28, 1967

B. Boehm (RAND)
G. Buck (UCSB)
L. Kleinrock (UCLA)
L. Roberts (ARPA)
B. Wessler (ARPA)

Enclosed are the first drafts of parts I (Introduction) and II (Scope) for the working group report on the IMP. Portions of the text and the appendices are left open where numbers, dates, etc. are still unknown. Section IIE (Network Operation and Maintenance) is unwritten due to uncertainties regarding the network operation following the installation and debugging stages.

Part III (Description of IMP Boundary Conditions) is now in the writing stage. In the process of attacking section IIB (IMP-Communication Facility Interface), I found it desirable to formulate a brief description of the communication functions. This description has been written separately, is now awaiting typing, and will hopefully be in the mail to you by December 4.

Very truly yours,

E. Shapiro

Elmer B. Shapiro
Senior Research Engineer

EBS:pc
Enc. 2

E. B. Shapiro
11/27/67

Draft Report of the ARPA Computer Network Working Group

I. Introduction

A. Network Objectives and Uses

A network of many interconnected computers is to be formed. These computers, called hosts, are or will be at the locations of ARPA contractors. Many such machines are now in use. The hosts are diverse in type, location, and application as indicated in the inventory of these machines provided in Appendix A, Inventory of Host Computers.

The network will consist of host computers, small stored program computers called interface message processors (IMP's), and communication transmission circuits and switches. This network is to permit the ready flow of digital information between host computers on a completely automatic basis. This information may be data or programs. Many of the host computers have consoles for use by people involved in many forms of research; time-sharing techniques are widely used to provide rapid console response, enabling a high degree man-machine interaction to occur. It is intended that the network will provide a vastly increased scope of services to be offered these users. The network responses are to be sufficiently fast to sustain the interactive process.

In addition to serving the host computers, the network itself is intended to be a subject of study and experimentation; for this reason, data gathering facilities will be incorporated into the network. A high degree of flexibility is to be provided, in part

through the use of a stored program computer for the IMP, to facilitate changes in the network (e.g., functions performed, techniques used, size).

B. Host-IMP Relationship

These host computers have input-output facilities, programs, storage capacity, computing capacity, and staff to serve their present or planned needs. It is ARPA's desire and intent to minimize the disruption, to these elements of the host system, caused by the advent of the network. The IMP's are intended to serve this minimization function by assuming, to a large extent, the additional burden imposed by the network on input-output facilities, programs, storage capacity, computing capacity, and staff. The functions to be performed by the IMP's are described in Section IV. These functions, in conjunction with the ARPA specified network conventions define the role of the IMP. In order to perform its job the IMP will consist of a basic hardware and a basic software package; these are to be delivered to the sites of specified host computers and there to be connected to the host computers and to the ARPA provided communication facilities. The managers of the host computers will provide (1) the hardware necessary to connect the host computer to the standard IMP-host interface, (2) host programs and perhaps some limited IMP programs to permit the host software system to work with the standard IMP software, and (3) the programs in the host to permit it to work with other hosts.

C. Host Types and Locations

The Inventory of Host Computers, Appendix A, lists 19 sites, containing XX computers--XX sites have more than one computer to be connected to the network. These computers represent X types, differentiating between manufacturers, and between models for a given manufacturer. These computers can serve XX terminals, some of which

are teletypewriter devices, and some of which are cathode-ray-tube devices with associated keyboards.

The host computers may, but need not, operate 24 hours per day, every day of the year. Typically, a host computer will be unavailable for communication with its IMP some interval during each day for purposes of preventative maintenance, emergency service, or experimentation. The unavailability of a host computer should not prevent its IMP from continuing to provide service to the remainder of the network.

D. Communication Features

High speed digital transmission, 50 kilobits per second (kbps) will be used extensively to enable fast network response times to be achieved; for a limited number of less demanding applications some 2.4 kbps circuits will be used. A tentative map of the communication facilities to be provided by ARPA, through DCA and the communication common carriers, is shown in Appendix B, Communication Circuit Map.

Point-to-point communications will be provided through the use of both non-switched and switched circuits for both the 50 kbps and 2.4 kbps transmission speeds. Each circuit will operate on a full duplex basis (for both switched and non-switched operation), with each direction at the indicated speed.

The routing of messages within the network will be controlled by IMP programs. These programs will use suggested ARPA routing doctrines that are sensitive to rapid changes (measured in seconds of time) in the characteristics of the generated traffic and to changes in the loading of the communication circuits. Other doctrines will be used to slowly (matter of minutes)

reconfigure the network by calling on the communication system, through dial-up and disconnect operations, to switch circuits in or out.

Control of errors arising in transmission and control of the rate of message flow will be vested in the IMP's, where programs will implement ARPA specified network conventions.

Appe : A

Inventory of Host Computers

Name of ARPA Contractor	Abbrev. of Name	Site Location	No. of Computers	Manufacturer and Model Number of Host Computer
Dartmouth College	DART	Hanover, N.H.		
	MIT			
	BBN			
	HARV			
	LL			
	BTL			
	ARPA			
	CMU			
	UM			
	UI			
	WU			
	UTAH			
	UCB			
Stanford Research Institute	SRI	333 Ravenswood Ave. Menlo Park, Calif. Bldg. 30, Rm. K-2079	2	SDS 940 SDS 940
	SU			
	UCSB			
	UCLA			
	RAND			
	SDC			

> Fenway Park, 9/27/55 Donald Davies.
Boston:

— (astonishment at the level of interest — seen as preposterous!)

— limited network:

— confirmed that he had given more thought to the interface; not much done w/redundancy →

— surprised by the level of support from Brit. tele/post.

— Bell official told Davies that packet switching wouldn't work.

— ~~B~~ saw packet model as most closely analogous to time-division ~~to~~ multiplexed

time-sharing = round-robin scheme.

Means of Providing a Short-Message Data Service

There has been some experience with digital transmission for P.C.M. speech over junctions. This shows that, using regenerators at the spacing normally used for loading coils, two pairs of wires can carry 1.5×10^6 bits/second. The reduction of cost as compared with the transmission of digital data over a speech path is about 25 : 1 (Hartley and Thomas, IEE Colloquium, October 1965).

The channel capacity provided is embarrassingly large. If 10^6 people in the London area sit at keyboards for an 8 hour day and send 100 calls per day on average, each of 50 characters, this amounts to 1.4×10^6 bits per second on average. Thus the capacity of a single P.C.M. junction would not be used on any one route.

Since there must be independent channels to allow for fault immunity, it seems likely that digital data will share the transmission equipment provided for speech. It will not necessarily influence speech towards the universal adoption of P.C.M., in fact data might always have to be carried less efficiently than is technically feasibl to suit the economics of speech transmission. Further speculation on the means of transmission is unnecessary since the cost of transmission can easily be seen to be very low.

The packing of the short messages onto a digital path at 10^6 bits per second is possible by short-term storage with very small delays. For smaller capacities, down to about 10^4 bits per second there will be no problem, but since delays may occur at each switching centre they must be individually no more than 100 milliseconds. This implies that 10^4 bits per second is about the lowest total transmission capacity between two switching centres which is allowable without partitioning the message into smaller pieces. Already at 50 characters the overhead due to routeing informatio will be relatively large and further partitioning seems undesirable.

A problem will arise in giving economical service in areas with a low concentrati of data stations, such as residential areas. Apart from this, the cost of transmissio will be very low, and it may happen that the capital cost of the computers handing the storage and multiplexing at each switching centre dominate the cost. This leads to a research topic: the system design and programming of message switching computers for the public network. Several examples of private systems exist, but there is, I believ no British firm in the business.

Other uses for a Short-Message Data Service

Such a system could take over the telegraph and telex services and make them more convenient because the other applications would cover the cost of many more sets of terminal equipment. The interaction with the system might have to be simplified so that anyone used to the keyboard could send messages as well as operate his own specialised services.

Ultimately the control of the switching system for telephones might use messages carried via the message network. This would be potentially cheaper and more reliable than using various forms of modulation on the telephone channels, and it would fit in with computer control. It would also allow more complicated automatic interactions with the telephone system such as personal calls set up via keyboards.

D. W. DAVIES

10th November, 1965

The big traffic/~~keyboard~~ for keyboard messages will be from professional operators. For example the staff dealing with the public in banks, post offices and Government offices will use them. Here the fixed format is not a difficulty. Other users will also be specialists in their job, such as designers using computational tools. People sending enquiries and placing orders for goods of all kinds will make up a large section of the traffic.

Suppose, however, that 10% of all the working population or 2% of the whole population makes frequent use of these keyboards, sending 100 messages a day each. This might compare with non-business use of the telephone amounting to several telephone calls per head per day. Business use of the telephone may be reduced by the growth of the kind of service we contemplate.

The overall result is that telephone calls and short data-messages will not be very different in number. Since any sensible engineering solution should pack a data message into the equivalent of a fraction of a second of a telephone channel, we can predict that the communications needs of data for on-line remote processing will be small compared with those of the telephone network.

We are assuming that the adoption of on-line data processing will largely remove the need for fast transmission of a considerable amount of data. Large quantities of data for transmission have either been accumulated over a period in the wrong place or generated by a computational process in the wrong place or generated by data-acquisition machinery. The sort of situation where a large experiment generates data which must be transmitted for reduction elsewhere can be assumed to be exceptional, and not necessarily to require the most economical provision for communication.

Qualities of the Service Needed

It is important to remark that, though its volume will be small, the short message traffic will be vital to the country. It will, in fact, carry more information than the telephone network. Corresponding to its much lower redundency, its requirement for noise-free transmission will be greater.

Each station sends and receives data at a low rate and sporadically. The unit message of about 50 characters probably takes 10 seconds to send to the centre and the reply takes 5 seconds to return. A delay of up to one second in each transmission might be permitted if it saved money.

When a station has connected with a certain data processing service, and the service centre has identified the station, subsequent messages should be passed through the communication system without further red-tape, but only occupying channel capacity when the keyboard is being used. Thus a keyboard could be permanently 'connected' to a distant service but only be occupying the minimum of local exchange equipment. The 'connection' would simply be an entry in a table in the controlling computer.

The accuracy worth striving for where humans generate and receive the data is limited, but one error in 10,000 characters would be a desirable goal. For those cases needing greater accuracy the system could provide for transmission by two paths and comparison, but the terminals would in that case need special attention in order not to add their own errors.

For security, the system might provide from a third source to each end of a transmission path a 'one time pad' on demand. This would ensure that nowhere in the network was the data in clear. There would, of course, be a limit to the degree of security that a public service could offer.

Remote on-line Data Processing and its Communication Needs

Remote on-line data processing is at a very early stage of practical achievement, so it is perhaps necessary to justify the making of long-term predictions, which are necessarily speculative. Long-term predictions may be valuable because they may indicate the future need for new kinds of communication service. The development of the telephone/telegraph system is constrained by present-day economics and by the inertia to change due to the need for all equipment to fit into the existing system. If new services are needed, or better ways of providing the present services can be forseen, requiring major changes, then planning well ahead is needed so that changes affecting the whole system can be started many years before they can be justified by immediate economics. It may also prove the research should be started now to determine how ~~but~~ best to provide new services in, say, ten years' time.

This note contains some very tentative predictions, intended to show the sort of considerations that might go into a more prolonged and detailed study.

Remote, On-line Data Processing

In this kind of data processing the computer system handles all the records in a well developed file-store and communicates at about 10 characters per second with people working at keyboards with simple printers and, eventually, cheap tabular and other displays. (The displays must have local storage to keep the communication rate economically feasible.) Since data is not accumulated at stations, faster transmission using paper tape cards or magnetic tape is not a requirement.

Larger users might have their own computers, which carry out simpler parts of the work, giving some immediate feedback to the users, and collect the data so that it can be transmitted for major processing jobs. This would need a higher transmission rate.

If the use of individual stations becomes economic, and this depends on the efficiency of communication, this kind of traffic will be greater than that between satellite computer and central service, because many more small organisations will be able to afford the simple equipment.

Already the few such services in existence (such as computation, stock exchange and airline reservation) show that they can be economical. Improvements in the economics of computers seem to be continuing, therefore the expansion of such services will not be limited by the cost. The security of the information is a problem, however, which leads to a subject of research: Security of data in doing business via a national network. (This refers to security against unauthorised access rather than against loss of data.)

Forecast of the Traffic

A typical transaction willconsist of about 50 characters sent at typewriting speed to the computing centre, and a reply of about the same size sent rather faster. These messages include identifying numbers and namos, the request, confirmation of the request and reply.

The greatest traffic could only come if the public used this means for everyday purposes such as shopping. It is doubtful whether the public can work accurately and confidently enough in a fixed format, and whether the £100 terminal cost (present-day prices with allowance for mass production) would be acceptable. There may be limited use of tone-button dials for simple transactions, but this needs no addition to the telephone service to provide it. Ultimately, looking further ahead, speech recognition and natural-language analysis might give the telephone system these new

EARLY PAPERS IN THE HISTORY OF PACKET SWITCHING

On Distributed Communications Networks

Paul Baran, IEEE Trans. on Communications Systems, March 1964

Remote on-line Data processing and its Communication Needs

Donald Davies, private communication, 10 November 1965

Futher Speculations on Data Transmission

Donald Davies, private communication, 16 November 1965

Proposal for the Development of a National Communication Service for on-line Data Processing

Donald Davies, private communication, 15 December 1965

Proposal for a Digital Communication Network

Donald Davies, private communication (widely circulated) June 1966

A Digital Communication Network for Computers Giving Rapid Response at Remote Terminals

Donald Davies, Keith Bartlett, Roger Scantlebury and Peter Wilkinson, ACM Symposium on Operating System Principles, Gatlinburg Tennessee, October 1967

Multiple Computer Networks and Intercomputer Communication

Lawrence Roberts, ACM Symosium, Gatlinburg, October 1967

Report on a Visit to the 1967 ACM Symposium USA

Roger Scantlebury, undated

Papers given at IFIP Congress 1968, Edinburgh

Communication Networks to Serve Rapid-Response Computers - Davies

The Principles of a Data Communication Network for Computers and Remote Terminals - Davies

The Design of a Message Switching Center for a Digital Communication Network - Scantlebury, Wikinson and Bartlett

The Control Functions in a Local Data Network - Wikinson and Scantlebury

Transmission Control in a Local Data Network - Bartlett

STANFORD RESEARCH INSTITUTE

MENLO PARK, CALIFORNIA 94025

December 27, 1967

B. Boehm (RAND)
G. Buck (UCSB)
L. Kleinrock (UCLA)
L. Roberts (ARPA)
B. Wessler (ARPA)

This letter confirms the arrangements for the next meeting of the ARPA Computer Network Working Group. This meeting is scheduled for 2 p.m., Tuesday, January 9, 1968, at the Stanford Research Instu tute, Menlo Park, California, Building 30, Room J2056. A map is enclosed to aid you in finding SRI, which is about 20 miles southeast of the San Francisco airport. From the airport one can travel by rented car (about a 30 minute drive), by taxicab (about $11 fare one-way), or by airport bus (an infrequent service that can get you to Palo Alto). For those driving the map is marked to show (1) the main entrance (marked on the map with a red 30) to Building 30, (2) the way into the preferred parking lot, and (3) the visitors' parking area in that lot (shown in shaded red). If that particular area is filled, you may occupy any vacant, no-reserved, parking space in any lot. Good second choice lots, in case of rainy weather, are those near the entrance marked on the map with the red 1. Once in SRI, please have a receptionist call me (Extension 2859) and you will be fetched.

Some suggested topics for the meeting are:

(1) a review of the work done to date,

(2) a reexamination of the time table, and

(3) development of a work plan for the immediate future.

We can draw up an agenda at the time we convene. Based on the previous meetings, we should plan for a four-hour session if things progress rapidly, a longer meeting otherwise.

Very truly yours,

E. B. Shapiro
Senior Research Engineer

EBS:pc
Enc.